

Quantity Perception by Adult Humans (*Homo sapiens*), Chimpanzees (*Pan troglodytes*), and Rhesus Macaques (*Macaca mulatta*) as a Function of Stimulus Organization

Michael J. Beran
Georgia State University, U.S.A.

Adult humans typically overestimate the number of items in regularly arranged stimulus sets compared to randomly arranged sets. This regular-random numerosity illusion (RRNI) was examined comparatively in adult humans, chimpanzees, and rhesus monkeys. Neither nonhuman primate species showed evidence of the illusion when trained to pick the larger of two sets of randomly arranged dots on a computer screen and then shown regularly arranged sets. Adult humans, given the same task and instructed to select the larger set, showed the illusion, although there were individual differences. Experiment 1 used somewhat different methodologies with the human participants compared to the nonhuman animals, but Experiment 2 presented the identical method to naïve human participants and naïve rhesus monkeys with minimized training, equal exposure to the different arrangement types, and very limited instructions to human participants. In that situation, human and monkey performance was very similar and reflected the RRNI. These results demonstrate that nonhuman animals also are susceptible to the RRNI, and they also indicate how methodological differences used during training both within and between species can impact results of comparative assessments.

Nonhuman animals are sensitive to numerical dimensions of stimulus sets. This sensitivity occurs with regard to enumerating the quantity of a single set of presented items (e.g., Biro & Matsuzawa, 2001; Boysen & Berntson, 1989; Hauser, MacNeilage, & Ware, 1996; Murofushi, 1997; Pepperberg, 1994), comparing two or more sets of items (e.g., Beran, 2001, 2004a, 2007; Beran & Beran, 2004; Brannon & Terrace, 2000; Call, 2000; Emmerton, Lohmann, & Niemann, 1997; Hauser, Carey, & Hauser, 2000; Judge, Evans, & Vyas, 2005; Olthof, Iden, & Roberts, 1997; Rumbaugh, Savage-Rumbaugh, & Hegel, 1987; Smith, Piel, & Candland, 2003; Thomas & Chase, 1980), and creating or enumerating sets of specified quantities (e.g., Beran, 2004b; Beran & Rumbaugh, 2001; Beran, Savage-Rumbaugh, & Rumbaugh, 1998; Davis & Bradford, 1986; Pepperberg, 1994; Roberts, Coughlin, & Roberts, 2000; Roberts, Roberts, & Kit, 2002; Rumbaugh, Hopkins, Washburn, & Savage-Rumbaugh, 1989; Xia, Emmerton, Siemann, & Delius, 2001; Xia, Siemann, & Delius, 2000). One area that has been overlooked, however, is how stimulus organization may operate to influence estimates of quantity in nonhuman animals. In fact, this area also has received little attention within the human literature. This is unfortunate given the interesting phenomena that emerge when one looks at how the presentation of stimulus sets leads to varying estimates of the number of items in those sets. For example, from their own observations of a

This research project was supported by Grant HD-38051 from the National Institute of Child Health and Human Development, Grant BCS-0634662 from the National Science Foundation and European Science Foundation, and the Rumbaugh Fellowship at Georgia State University. The author thanks John Kelley, Natasha Barrett, and Frances James for assistance with data collection and David Washburn for his helpful comments on an earlier version of this paper. Correspondence concerning this article should be addressed to Michael J. Beran, Language Research Center, Georgia State University, Atlanta, GA 30303. (mjberan@yahoo.com).

solitaire marble board, Frith and Frith (1972) discovered that certain configurations gave the impression that one colored set of marbles was more numerous than the other, when in fact both were equal in number. They found that this illusion occurred for children and adults whenever one large cluster was compared to several smaller clusters, and they argued that the illusion occurred because the number of items in the smaller clusters was underestimated (also see Ginsburg, 1982). Ginsburg and Nicholls (1988) reported that undergraduates tended to underestimate set sizes when stimuli (dots) were larger in diameter compared to estimates for the same number of dots with smaller diameters (but see Allik, Tuulmets, & Vos, 1991).

Ginsburg (1976) presented adult human participants with two sets of dots arranged in a circular pattern and asked the participants to choose the set that was more numerous. One set was arranged in a regular pattern whereas the other set was arranged randomly. When both sets contained the same number of items, participants' selections indicated that the regular arrangements appeared more numerous (also see Ginsburg, 1978). The use of a rectangular area of presentation like that used in the present study also led to the same trends for overestimation of regularly arranged sets and underestimation of randomly arranged sets. Ginsburg (1980) called this the *regular-random numerosity illusion* (RRNI).

The RRNI emerges when homogeneous sets of stimuli are presented (e.g., identical dots presented in certain configurations). In fact, that is one of its primary features, the use of homogeneous stimuli. Within the comparative literature examining numerical cognition of nonhuman animals, this use of homogeneous stimuli can be a problem if one wants to argue that judgments or responses are made on the basis of the numerosity of a set. This is because, with homogeneous sets, other stimulus dimensions co-vary with number perfectly, such as amount or area. For this reason, researchers strive to control for these non-numerical properties of the stimuli that they use, to provide compelling evidence that number is the controlling dimension (see Beran, 2007; Brannon & Terrace, 2000; Cantlon & Brannon, 2006; Judge et al., 2005; Smith et al., 2003). In the case of the RRNI, however, no such controls are used because the illusion itself is based on homogeneity of the stimuli.

To date, only humans have been tested for the RRNI, but a comparative assessment of quantity judgments by three species, adult humans, chimpanzees, and rhesus monkeys now is presented. The nonhuman primate data are presented to examine the extent to which the RRNI is shown across species, and what similarities or differences in this illusion across species might indicate about quantity perception. In addition, training parameters with nonhuman primates are varied to examine the extent to which training requirements affect subsequent response biases relevant to the RRNI.

In the following experiments, adult humans and two species of nonhuman primate, chimpanzees (*Pan troglodytes*) and rhesus monkeys (*Macaca mulatta*), were presented with a quantity judgment task in which the larger of two sets of homogeneous dots had to be selected. The majority of trials had a correct response with one set being larger than the other. Probe trials, however, were interspersed in which both sets had equal numbers of identical dots, but in which one set consisted of regularly arranged dots and the other set consisted of randomly arranged dots. If human participants showed a bias toward selecting the regularly arranged arrays

across differing quantities, this would replicate the findings of Ginsburg (1976, 1978, 1980). In addition, if the nonhuman primates showed similar biases, this would indicate a predisposition on the part of these species also to overestimate the quantity of regularly arranged items in comparison with randomly arranged items.

The purpose of this study is not to demonstrate that these nonhuman primate species can use number as the discriminative cue in a judgment task. This has been demonstrated using other computerized tasks in which controls were used for preventing stimulus distance, surface area, and other non-numerical aspects of stimuli from acting as valid cues for responding (e.g., Beran, 2007; Brannon & Terrace, 2000; Judge et al., 2005), and such controls also could be used in this task to the end of producing responding based on number. However, the concern of this study is the presence or absence of the RRNI in the judgments of nonhuman primates and humans when stimulus sets do not explicitly control for nonnumerical cues.

Experiment 1

Method

Participants. Human participants were recruited from Georgia State University, a large urban university in the southeastern United States. All participants received course credit for their participation. Seventeen participants completed the study (mean age, 19.88 years, range 18-22 years). There were 5 females included in the sample, and 7 of the participants were minority students.

All nonhuman primates came from the Language Research Center of Georgia State University. Four chimpanzees, Lana (female, 34 years of age), Sherman (male, 31 years of age), Panzee (female, 19 years of age), and Mercury (male, 18 years of age) were tested. Five male rhesus monkeys Murph, (11 years of age), Lou (11 years of age), Gale (21 years of age), Hank (21 years of age), and Willie (19 years of age) were tested. These animals all had extensive testing experience using joysticks to respond to computer-generated stimuli (see Richardson, Washburn, Hopkins, Savage-Rumbaugh, & Rumbaugh, 1990; Rumbaugh, Richardson, Washburn, Savage-Rumbaugh, & Hopkins, 1989), including tests relevant to areas of numerical cognition (e.g., Beran, 2004b; 2007; Beran & Rumbaugh, 2001; Harris & Washburn, 2005; Washburn & Rumbaugh, 1991).

Apparatus. All participants completed trials that were presented on a 17 inch CRT computer monitor attached to a personal computer running Microsoft Windows. The experimental program was written in Visual Basic 6.0 for Windows. Human participants sat at a table approximately 1 m from the monitor and responded through keyboard presses. All nonhuman animals were observed individually in their home cages. Joysticks were mounted on the cages so that the animals could manipulate the joysticks with their hands, and all animals were thoroughly familiar with the use of these joysticks. The monkeys viewed the monitor through cage wire from a distance of approximately 45 cm, and the joystick was located approximately 15 cm below the monitor. The chimpanzees viewed the monitor through cage wire from a distance of approximately 75 cm, and the joystick was located approximately 60 cm below the monitor.

Design and Procedure. Human participants were instructed to contact the space bar to begin each trial. After contacting the space bar, participants saw two sets of circular black dots (4.5 mm in diameter) on a white background on the monitor, and they were instructed to indicate which of the two sets was larger. If the left set was larger, they were instructed to press the "A" button. If the right set was larger, they were instructed to press the "L" button. Participants had to make one of these responses within 1.5 s for the trial to count toward the 250 trials that constituted a test session. If participants did not make a response within that time constraint, the trial was terminated (and did not count toward the 250 trials), and a note was presented on the screen asking the participant to respond more quickly. This requirement for fast responses was necessary to make sure participants were not able to count the items in either set. Additionally, this time constraint prevented participants from

counting the number of items in each row and column and using multiplication to assess set size when sets were arranged in a regular pattern. Participants did not receive any feedback as to whether their responses were correct or incorrect. To encourage participants to perform at high levels, a cash prize of \$10 was offered to the highest scoring participant.

The majority of trials (85%) involved two sets of randomly arranged dots in which one set was larger than the other. For those trials, each set could contain from 1 to 40 dots (4.5 mm in diameter), and set sizes were determined randomly by the program with the constraint that the larger set was not more than 20 items larger than the smaller set. Each dot was randomly assigned to 1 of 121 locations within an 11x11 grid on each side of the computer screen. The program randomly determined whether the larger set was presented on the left half of the monitor or the right half of the monitor. The remaining 15% of the trials involved two sets equal in number. These equal sets could contain 12, 16, 20, 25, 30, or 36 dots (see Figure 1). On one side of the screen (randomly determined), the dots were placed in a regular arrangement. If the set size was 12, this was an arrangement of 4 rows with 3 items in each row, and each dot was equally spaced from all dots surrounding it. If the set size was 16, this was an arrangement of 4 rows with 4 items in each row. If the set size was 20, this was an arrangement of 4 rows with 5 items in each row. If the set size was 25, this was an arrangement of 5 rows with 5 items in each row. If the set size was 30, this was an arrangement of 5 rows with 6 items in each row. If the set size was 36, this was an arrangement of 6 rows with 6 items in each row.

Because some of these regular arrangements covered a smaller total area on the screen, the randomly arranged comparison set also was arranged across a similarly smaller area compared to the baseline trials in which set sizes were different. If the number of dots in the regularly arranged set was 12 or 16 items, the randomly arranged set was placed within a 10x10 grid on the screen. If the number of dots in the regularly arranged set was 20, 25 or 30 items, the randomly arranged set was placed within a 9x10 grid on the screen. If the number of dots in the regularly arranged set was 36 items, the randomly arranged set was placed within the regular 11x11 grid on the screen. This ensured that both sets covered approximately the same total area on the computer screen. Participants each completed 250 trials in which they made the primary response before the 1.5 s interval had ended. Although participants occasionally failed to respond quickly enough, this was rare (less than 1.5% of the trials).

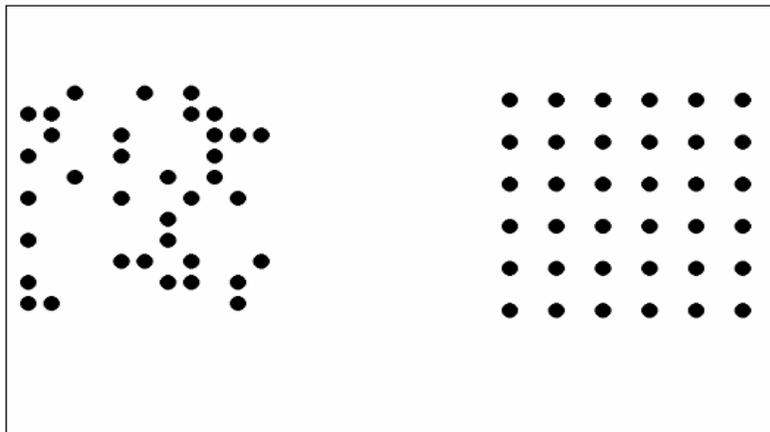


Figure 1. An example probe trial. The left set contains the random arrangement, and the right set contains the regular arrangement. Both sets contain 36 items. However, without counting the sets, many people estimate that the set on the right has more items, and that is the RRNI.

The procedure for the nonhuman primates was modified. First, the animals responded through use of a manually controlled joystick that moved a cursor into contact with one of the two sets of dots presented on the screen, and each set contained from 1 to 40 items. Second, nonhuman primates had to be trained as to the task demands, and this training took the form of presenting baseline trials until an animal reached a criterion of 18 trials correct of the last 20 trials performed. These

baseline trials consisted of the presentation of two randomly arranged sets. Baseline trials were presented at the start of all test sessions, and this same criterion was used for all sessions with the nonhuman animals. Third, nonhuman primates were rewarded for correct responses on all baseline trials. The monkeys received 94-mg banana flavored food pellets that were dispensed automatically by a feeder attached to the computer apparatus. Incorrect responses led to a 10 s timeout during which the screen was blank. Chimpanzees received small bits of preferred food items (e.g., banana slices, grapes) for correct responses to these trials. Incorrect responses led to no food reward. Fourth, there was no response time criterion for the nonhuman primates.

After reaching criterion, nonhuman primates then shifted to the 85%-15% division of baseline and probe trials identical to those given to the human participants. Four of the nonhuman primates received no positive or negative feedback on probe trials (monkeys Lou and Gale and chimpanzees Lana and Sherman). For these animals, a new trial was presented immediately after a response to a probe trial. The other five animals were rewarded for either selection on probe trials.

Each animal participated in multiple testing sessions to accumulate a large number of probe trials. Murph completed a total of 2,209 trials, including 339 probe trials, across two test sessions. Lou completed a total of 2,292 trials, including 355 probe trials, across three test sessions. Gale completed a total of 4,437 trials, including 639 probe trials, across three test sessions. Hank completed a total of 3,861 trials, including 542 probe trials, across three test sessions. Willie completed a total of 4,964 trials, including 679 probe trials, across three test sessions. Each chimpanzee completed a total of 1,000 trials across seven test sessions. Respectively, Lana, Mercury, Panzee, and Sherman completed 135, 128, 147, and 108 probe trials within the 1000 total trials.

Individual responses to probe trials were combined across all human participants for subsequent analyses. For the nonhuman primates, data were combined within each species. Analyses of probe trials were conducted using the data from individual trials (i.e., frequency data) because individual human participants did not complete a sufficiently large number of probe trials at each set size for their mean percentages of random or regular selections to be useful, and there were not a sufficient number of participants for each nonhuman primate species to use mean scores. Where reported in isolation, p values are the result of analyses using a two-tailed sign test because chance was estimated to be 50% and because I was interested in response preferences for random or regular arrangements of stimulus sets.

Results

Prior to examining selection biases on the probe trials, I first examined overall performance of all groups in the baseline quantity judgment task. There was little difference in the performance of nonhuman primates receiving feedback after probe trials (chimpanzees, 79.4%; monkeys, 81.6%) and those not receiving feedback (chimpanzees, 74.1%; monkeys, 86.0%). Thus, data were combined across individuals within each species for subsequent analyses. All three species showed reliable distance effects as performance increased with larger quantitative differences between the two sets of dots (humans, $r(18) = .88$, $p < .01$; chimpanzees $r(36) = .86$, $p < .01$; monkeys, $r(37) = .89$, $p < .01$; Figure 2).

The selection patterns of all groups on the probe trials are illustrated in Figure 3. Adult humans were the only group to show the RRNI, as they selected the regularly arranged set on significantly more trials than expected by chance for sets of 20, 25, 30, and 36 items, $p < .05$. For the nonhuman primates, only two set sizes showed distributions of responses that differed from chance on probe trials, and in both cases the animals selected the randomly arranged set more often than expected by chance. Chimpanzees selected the randomly arranged set more often for sets of 12 items, $p < .05$. Monkeys selected the randomly arranged set more often for sets of 16 items, $p < .05$.

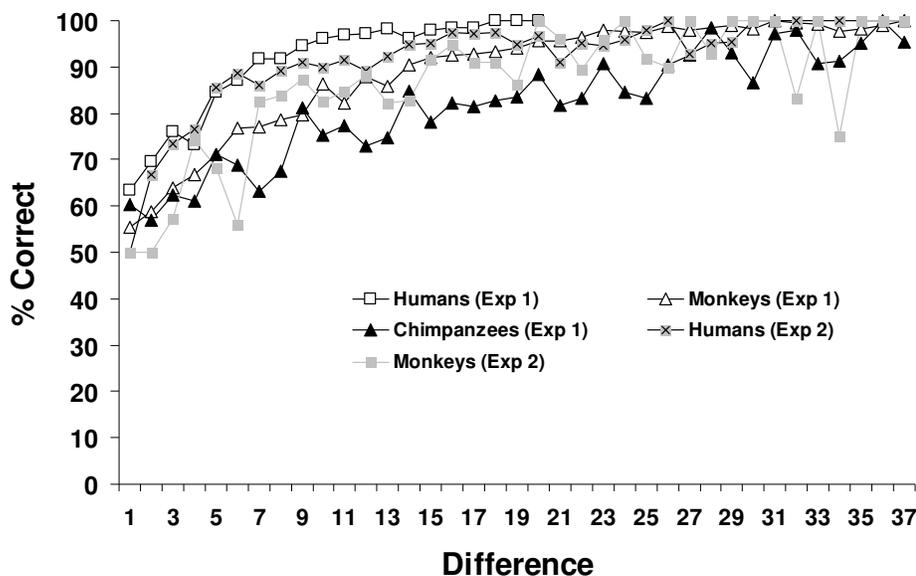


Figure 2. Performance of each species on baseline trials as a function of the quantitative difference between sets. All three species showed a distance effect in that performance increased as the difference between sets increased. Human and monkey data are shown for Experiment 1 and Experiment 2 (chimpanzees were not tested in Experiment 2).

I also examined individual differences within species on this task by collapsing the data across all set sizes (thus ensuring conservative estimates of biases for one type of arrangement over the other given that none of the species showed consistent biases across all set sizes). Two chimpanzees showed an overall statistically significant bias toward selecting the randomly arranged set on probe trials (Lana and Mercury, $p < .05$), but the other two chimpanzees showed no such biases overall. Two of the five monkeys (Willie and Lou) also showed an overall statistically significant bias toward selecting the randomly arranged set on probe trials ($p < .05$), but the other three monkeys showed no such biases overall. Among adult humans, 14 of the 17 participants showed biases that were statistically different from chance, $p < .05$. However, 5 of the 17 participants actually had a bias toward selecting the randomly arranged set as being the larger set whereas 9 participants had a bias toward the regularly arranged set (leaving 3 participants with no bias toward either set).

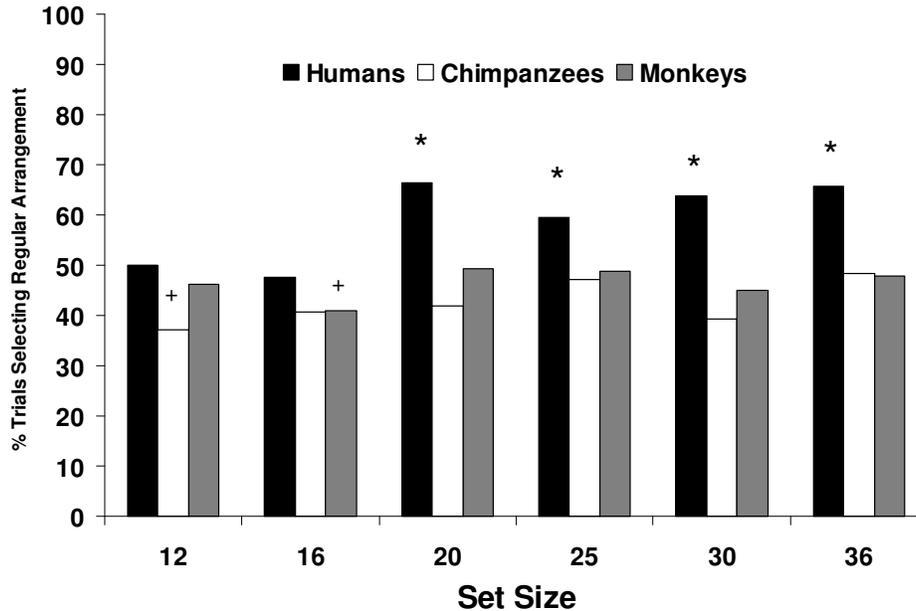


Figure 3. Arrangement biases for each species as a function of set size. Only humans showed a statistically significant preference for the regularly arranged set sizes ($p < .05$, indicated by the * in the figure). In two cases, nonhuman primates showed a statistically significant preference for the randomly arranged set ($p < .05$, indicated by the + in the figure).

Discussion

Three results emerged from this study. First, when selecting the larger of two sets of items, adult humans as a group estimated regularly arranged sets to be more numerous than randomly arranged sets of items. They did this when asked to respond quickly and without the opportunity to count the numbers of items in each set. It should be noted, however, that this overestimation did not occur for the smallest set sizes. The most likely explanation for this is that the response time criterion was too long, and participants may have been given too much time to use strategies such as multiplication or rapid enumeration. The finding of overestimation for the larger set sizes replicates earlier studies showing a bias to perceive regularly arranged sets as containing more items than randomly arranged sets (e.g., Ginsburg, 1976, 1980).

The second result was less expected. Some human adults actually showed a bias toward randomly arranged sets of stimulus when asked to pick the larger set. Thus, an examination of individual differences, which was not often undertaken in previous studies, indicates that the RRNI may not hold across all individuals, and that some people actually may exhibit the opposite bias toward randomly arranged sets.

The third result is that the two species of nonhuman primates showed no such perceptual illusion in overestimating regularly arranged sets compared to randomly arranged sets. In fact, if any trend was present, it was a trend toward per-

ceiving randomly arranged sets as being more numerous in these species as all individual animals that showed biases did so for the randomly arranged sets. Thus, there appears to be a discontinuity across species with regard to this illusion in estimating quantity. However, this difference may have been the result of the different training requirements used with the nonhuman primates. Because the nonhuman primates received baseline trials, all of which involved randomly arranged sets, they may have simply avoided choosing the regularly arranged arrays because they had experienced fewer rewards for selecting stimuli in these arrangements. Thus, when sets were viewed as being very similar in quantity, the nonhuman primates may have selected the more familiar arrangements. This would have led to a bias toward the randomly arranged sets, which is exactly what was found for some of the nonhuman primates. Therefore, a second experiment was conducted in which adult humans and rhesus monkeys received the identical computerized task, and nearly all details of the testing situation were equated. This allowed an examination of whether the failure to find the RRNI in the nonhuman animals was due to methodological differences or actual perceptual differences between these two species.

Experiment 2

Method

Participants. Twenty-four additional participants from Georgia State University were tested (mean age, 20.75 years, range 18-41 years). There were 21 females included in the sample, and 15 of the participants were minority students. All participants received course credit for their participation. Four new male rhesus monkeys (Luke, Han, Obi, and Chewie) were tested in this experiment. They ranged in age from 3 years to 7 years. None of these monkeys had ever participated in quantity judgment tasks prior to the start of the experiment. No additional chimpanzees were tested because all available animals had participated in Experiment 1.

Design and Procedure. Both species completed the same task. All participants completed 300 trials in this experiment, not counting trials excluded because responses were not made within 1.5 s. On each trial, a participant moved the cursor into contact with a start button in the center of the screen by using a joystick. Two arrays of dots then appeared, and the participant had 1.5 s to make a response. During training, there were equal numbers of trials presented in which the participant was shown two randomly arranged arrays, two regularly arranged arrays, or one array of each type. Therefore, in this experiment, familiarity with both arrangement types was equated during training. Each participant was trained to a criterion of 16 of the last 20 trials correct before trials could appear with equal numbers of dots (one set arranged randomly and one arranged regularly exactly as in Experiment 1). These probe trials appeared on approximately 30% of the trials after criterion was met, and no feedback was given for either response. Feedback (pellets for monkeys and the word "Correct" for humans) was given for all trials in which the number of dots differed between the two sets and a correct response was made. Incorrect responses to unequal sets led to a 20 s timeout for monkeys or a 5 s timeout for humans. For both species, if a response was not made within 1.5 s the message "Too Slow" appeared, and then the penalty timeout began. This occurred only rarely (less than 1% of the trials for humans and less than 3% of the trials for monkeys).

The only difference between the monkey procedure and the human procedure pertained to the instructions that humans received and the form of feedback they received. Human participants saw the following instructions: "Thank you for your participation. Please use the joystick in front of you to move the cursor (the red dot) on the screen. Start each trial by touching the rectangle in the center of the screen. Then, you will see two sets of dots. For the first part of the experiment, after you pick one set you will be told if your choice was correct or not, and you should try to learn the simple rule for making choices. Later, there may be some trials with no feedback given and others where you

are told if you are correct or not, but the rule will still be the same so keep responding the same way. You will need to respond quickly in this task (within 1.5 seconds on each trial). If you do not, an error message will appear. You will do 300 trials where you make a response in time. The number of trials remaining is shown in the top left corner. Please try to get as many trials correct as you can using the simple rule that you will learn at the beginning.”

Results

All trials were again grouped for each species to provide a large corpus of data. For baseline trials, there was again a strong correlation of quantitative distance between sets and performance for humans, $r(37) = .76, p < .01$, and for the monkeys, $r(35) = .73, p < .01$ (Figure 2). For all 6 set sizes, human participants showed a statistically significant bias in choosing the regular arrangement over the random arrangement, $p < .05$. Additionally, the monkeys showed a statistically significant bias in choosing the regular arrangement over the random arrangement, $p < .05$, for all set sizes except for sets of 12 items (Figure 4).

An examination of individual differences again revealed that, although the majority of human participants (16 of 24) showed a statistically significant bias ($p < .05$), not all participants were biased to select the regularly arranged sets. Two of the 16 human participants who showed a bias chose the randomly arranged sets more often. At an individual level, 3 of 4 monkeys showed a significant bias ($p < .05$), and for all three that bias was to select the regularly arranged sets.

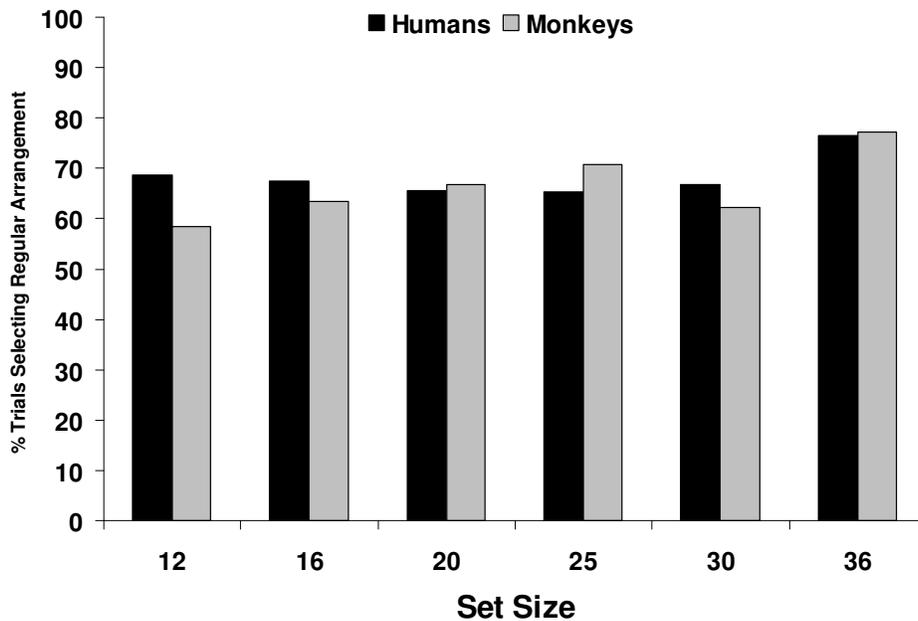


Figure 4. Arrangement biases for each species as a function of set size. Humans and monkeys showed a statistically significant preference for the regularly arranged set for all set sizes ($p < .05$) with the exception of the set size of 12 for the monkeys.

General Discussion

In Experiment 1, adult human participants showed a bias toward perceiving regularly arranged sets of stimuli as containing the larger quantity of items compared to randomly arranged sets of stimuli even though both sets had the same number of dots. This finding matched previous studies with adult humans (e.g., Ginsburg, 1976, 1978). However, a closer look at individual differences revealed that not all participants were susceptible to this illusion. In fact, some showed the opposite bias. Although one would expect that in any study there would be a minority of participants who do not follow the overall trend of the group, this issue must be remembered when discussing illusions in number judgments as a function of stimulus organization in adult humans.

Experiment 1 also assessed whether two nonhuman primate species, chimpanzees and rhesus monkeys, experienced the RRNI. The data from that experiment suggested that they did not. In fact, the nonhuman primates were more biased toward the randomly arranged sets than were human participants. Thus, the arrangements of stimuli produced clear differences in performance between humans and nonhuman primates. However, methodological changes between human and nonhuman testing may, in fact, have led to data that were not indicative of the true susceptibility of nonhuman animals to this illusion. Because the nonhuman primates were trained using random arrays of stimuli, it was possible their lack of bias to choose the regular arrays was an artifact of that training rather than a difference in perceptual processing.

The data from Experiment 2 confirmed this possibility. Four new rhesus monkeys looked very similar to humans when the methods were nearly identical across species. Therefore, the RRNI is a shared illusion, and one that suggests that nonhuman animals, like humans, may organize quantitative sets in such a way that regularly arranged sets somehow appear to contain more items. One possible explanation for the RRNI is based upon a Gestalt effect (Frith & Frith, 1972; Ginsburg, 1980) through which a single cluster of items constitutes a better gestalt than multiple clusters of items. Elements that form a better gestalt should appear more numerous, and good patterns are those with consistency (i.e., regularity) across the figure (see Ginsburg, 1980; Ginsburg & Nicholls, 1988). We know far less about the impact of Gestalt-like principles of organization on nonhuman primate perception than on human perception (but see Fujita, 1997, 2001), but these data indicate that perceptual “goodness” of stimulus organization may operate at comparable levels across species.

Finally, the difference between monkey performance in Experiment 1 and Experiment 2 illustrates an important point for comparative assessments of perception and cognition when one includes data from nonhuman animals and humans. Tasks used with humans often must be modified for use with nonhuman animals because of the inability to provide verbal instructions. These modifications may entail extended training that in some way leads to differing performance levels not indicative of true cognitive or perceptual differences but rather performance levels that reflect different training experiences. This should be kept in mind during task development for comparative assessments.

References

- Allik, J., Tuulmets, T., & Vos, P. G. (1991). Size invariance in visual number discrimination. *Psychological Research*, **53**, 290-295.
- Beran, M. J. (2001). Summation and numerosness judgments of sequentially presented sets of items by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, **115**, 181-191.
- Beran, M. J. (2004a). Chimpanzees (*Pan troglodytes*) respond to nonvisible sets after one-by-one addition and removal of items. *Journal of Comparative Psychology*, **118**, 25-36.
- Beran, M. J. (2004b). Long-term retention of the differential values of Arabic numerals by chimpanzees (*Pan troglodytes*). *Animal Cognition*, **7**, 86-92.
- Beran, M. J. (2007). Rhesus monkeys (*Macaca mulatta*) enumerate sequentially presented sets of items using analog numerical representations. *Journal of Experimental Psychology: Animal Behavior Processes*, **33**, 42-54.
- Beran, M. J., & Beran, M. M. (2004). Chimpanzees remember the results of one-by-one addition of food items to sets over extended time periods. *Psychological Science*, **15**, 94-99.
- Beran, M. J., & Rumbaugh, D. M. (2001). "Constructive" enumeration by chimpanzees (*Pan troglodytes*) on a computerized task. *Animal Cognition*, **4**, 81-89.
- Beran, M. J., Rumbaugh, D. M., & Savage-Rumbaugh, E. S. (1998). Chimpanzee (*Pan troglodytes*) counting in a computerized testing paradigm. *Psychological Record*, **48**, 3-19.
- Biro, D., & Matsuzawa, T. (2001). Use of numerical symbols by the chimpanzee (*Pan troglodytes*): Cardinals, ordinals, and the introduction of zero. *Animal Cognition*, **4**, 193-199.
- Boysen, S. T., & Berntson, G. G. (1989). Numerical competence in a chimpanzee (*Pan troglodytes*). *Journal of Comparative Psychology*, **103**, 23-31.
- Brannon, E. M., & Terrace, H. S. (2000). Representation of the numerosities 1-9 by rhesus macaques (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes*, **26**, 31-49.
- Call, J. (2000). Estimating and operating on discrete quantities in orangutans (*Pongo pygmaeus*). *Journal of Comparative Psychology*, **114**, 136-147.
- Cantlon, J. F., & Brannon, E. M. (2006). Shared system for ordering small and large numbers in monkeys and humans. *Psychological Science*, **17**, 401-406.
- Davis, H., & Bradford, S. A. (1986). Counting behavior by rats in a simulated natural environment. *Ethology*, **73**, 265-280.
- Emmerton, J., Lohmann, A., & Niemann, J. (1997). Pigeons' serial ordering of numerosity with visual arrays. *Animal Learning and Behavior*, **25**, 234-244.
- Frith, C. D., & Frith, U. (1972). The solitary illusion: An illusion of numerosity. *Perception and Psychophysics*, **11**, 409-410.
- Fujita, K. (1997). Perception of the Ponzo illusion by rhesus monkeys, chimpanzees, and humans: Similarity and difference in the three primate species. *Perception and Psychophysics*, **59**, 284-292.
- Fujita, K. (2001). Perceptual completion in rhesus monkeys (*Macaca mulatta*) and pigeons (*Columba livia*). *Perception and Psychophysics*, **63**, 115-125.
- Ginsburg, N. (1976). Effect of item arrangement on perceived numerosity: Randomness vs regularity. *Perceptual and Motor Skills*, **43**, 663-668.
- Ginsburg, N. (1978). Perceived numerosity, item arrangement, and expectancy. *American Journal of Psychology*, **91**, 267-273.
- Ginsburg, N. (1980). The regular-random numerosity illusion: Rectangular patterns. *Journal of General Psychology*, **103**, 211-216.

- Ginsburg, N. (1982). Solitaire illusion: Replication and extension. *Perceptual and Motor Skills*, **54**, 209-210.
- Ginsburg, N., & Nicholls, A. (1988). Perceived numerosity as a function of item size. *Perceptual and Motor Skills*, **67**, 656-658.
- Harris, E. H., & Washburn, D. A. (2005). Macaques' (*Macaca mulatta*) use of numerical cues in maze trials. *Animal Cognition*, **8**, 190-199.
- Hauser, M. D., Carey, S., & Hauser, L. B. (2000). Spontaneous number representation in semi-free-ranging rhesus monkeys. *Proceedings of the Royal Society of London B*, **267**, 829-833.
- Hauser, M. D., MacNeilage, P., & Ware, M. (1996). Numerical representation in primates. *Proceedings of the National Academy of Sciences*, **93**, 1514-1517.
- Judge, P. G., Evans, T. A., & Vyas, D. K. (2005). Ordinal representation of numeric quantities by brown capuchin monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behavior Processes*, **31**, 79-94.
- Murofushi, K. (1997). Numerical matching behavior by a chimpanzee (*Pan troglodytes*): Subitizing and analogue magnitude estimation. *Japanese Psychological Research*, **39**, 140-153.
- Olthof, A., Iden, C. M., & Roberts, W. A. (1997). Judgments of ordinality and summation of number symbols by squirrel monkeys (*Saimiri sciureus*). *Journal of Experimental Psychology: Animal Behavior Processes*, **23**, 325-339.
- Pepperberg, I. M. (1994). Numerical competence in an African Grey parrot (*Psittacus erithacus*). *Journal of Comparative Psychology*, **108**, 36-44.
- Richardson, W. K., Washburn, D. A., Hopkins, W. D., Savage-Rumbaugh, E. S., & Rumbaugh, D. M. (1990). The NASA/LRC Computerized Test System. *Behavior Research Methods, Instruments, and Computers*, **22**, 127-131.
- Roberts, W. A., Coughlin, R., & Roberts, S. (2000). Pigeons flexibly time or count on cue. *Psychological Science*, **11**, 218-222.
- Roberts, W. A., Roberts, S., & Kit, K. A. (2002). Pigeons presented with sequences of light flashes use behavior to count but not to time. *Journal of Experimental Psychology: Animal Behavior Processes*, **28**, 137-150.
- Rumbaugh, D. M., Hopkins, W. D., Washburn, D. A., & Savage-Rumbaugh, E. S. (1989). Lana chimpanzee learns to count by "NUMATH": A summary of a videotaped experimental report. *Psychological Record*, **39**, 459-470.
- Rumbaugh, D. M., Richardson, W. K., Washburn, D. A., Savage-Rumbaugh, E. S., & Hopkins, W. D. (1989). Rhesus monkeys (*Macaca mulatta*), video tasks, and implications for stimulus-response spatial contiguity. *Journal of Comparative Psychology*, **103**, 32-38.
- Rumbaugh, D. M., Savage-Rumbaugh, E. S., & Hegel, M. T. (1987). Summation in the chimpanzee (*Pan troglodytes*). *Journal of Experimental Psychology: Animal Behavior Processes*, **13**, 107-115.
- Smith, B. R., Piel, A. K., & Candland, D. K. (2003). Numerity of a socially housed hamadryas baboon (*Papio hamadryas*) and a socially housed squirrel monkey (*Saimiri sciureus*). *Journal of Comparative Psychology*, **117**, 217-225.
- Thomas, R. K., & Chase, L. (1980). Relative numerosity judgments by squirrel monkeys. *Bulletin of the Psychonomic Society*, **16**, 79-82.
- Washburn, D. A., & Rumbaugh, D. M. (1991). Ordinal judgments of numerical symbols by macaques (*Macaca mulatta*). *Psychological Science*, **2**, 190-193.
- Xia, L., Emmerton, J., Siemann, M., & Delius, J. D. (2001). Pigeons (*Columba livia*) learn to link numerosities with symbols. *Journal of Comparative Psychology*, **115**, 83-91.
- Xia, L., Siemann, M., & Delius, J. D. (2000). Matching of numerical symbols with number of responses by pigeons. *Animal Cognition*, **3**, 35-43.